Stat 145, Fri 24-Sep-2021 -- Fri 24-Sep-2021
Biostatistics
Spring 2021

Friday, September 24th 2021

Wk 4, Fr
Topic:: 95 percent confidence intervals
Topic:: Estimating SE using bootstrapping
Read:: Lock5 3.3

CI construction: 95%
- goal: to estimate population parameter
    frequently: mu, p, mu_1 - mu_2, p_1 - p_2
    Why? We already have unbiased estimators (sample statistics)
        mu: x-bar
        p: p-hat
        mu_1 - mu_2: x-bar_1 - x-bar_2
        p_1 - p_2: p-hat_1 - p-hat_2

- How:
    1. centered interval approach
        take estimate +/- (2)(SE)
        2*SE is called the margin of error (specific to 95% confidence)

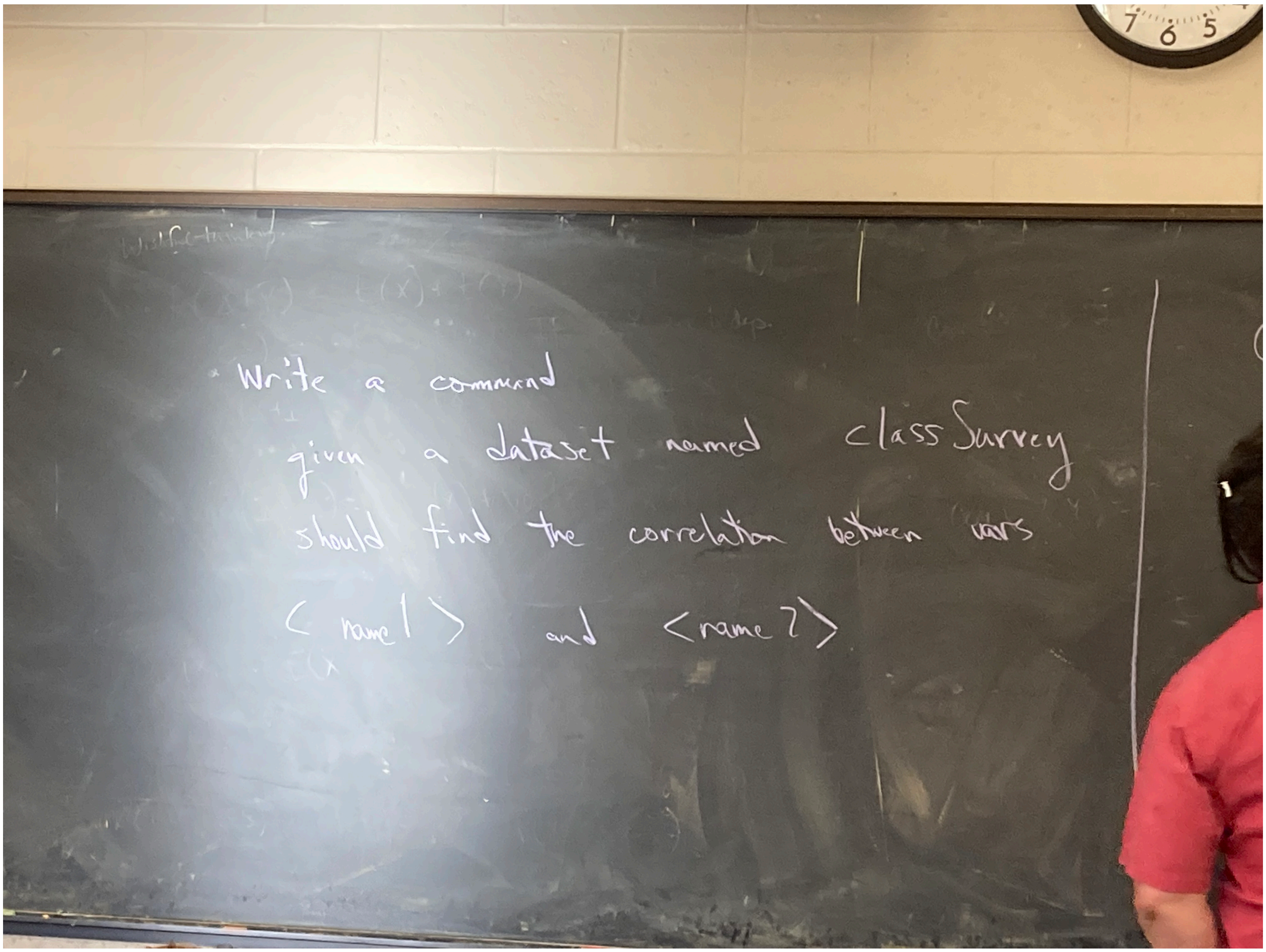
    2. percentile approach (must await discussion of bootstrapping)

- Meaning of CI
    Note the three misinterpretations the Locks want you to avoid, pp. 187-88

- Example:

Belief	Females	Males
There isn't one true love	1005	807
There is one true love	363	372
Undecided	34	44

If SE = .018, find 95% CI for difference in proportion who disagree



Write a command
given a dataset named classSurvey
should find the correlation between vars
< name1 > and <name2>

Confidence Intervals

Goal: Estimate

- σ
- μ
- p (pop. proportion)
- ρ (pop. correlation)
- β_1 (pop. slope)
- $\mu_1 - \mu_2$
- $p_1 - p_2$

Always off

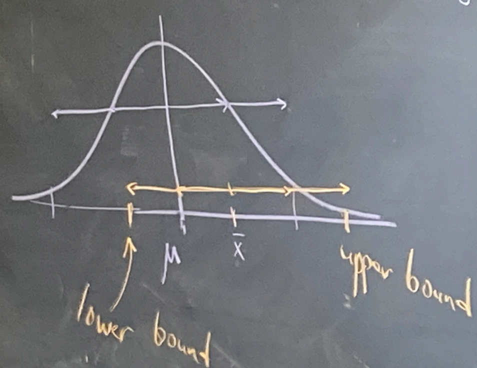
<u>Have estimators</u> (sample stats)

- s
- $\bar{x}$
- $\hat{p}$
- r
- b_1
- $\bar{x}_1 - \bar{x}_2$
- $\hat{p}_1 - \hat{p}_2$

Want μ

Have $\bar{x}$ (sampling dist normal)

95% of values are within $2 \cdot SE$'s of center



Our 1st method for contructing 95% CI:

$$(\text{point est}) \pm 2 \cdot SE$$

Defn: The amount added/subtracted from estimate to obtain lower/upper bounds of your CI is called the **margin of error**.

Ex.) Have poll saying 42% of Americans favor Proposal A. (This is $\hat{p}$.)

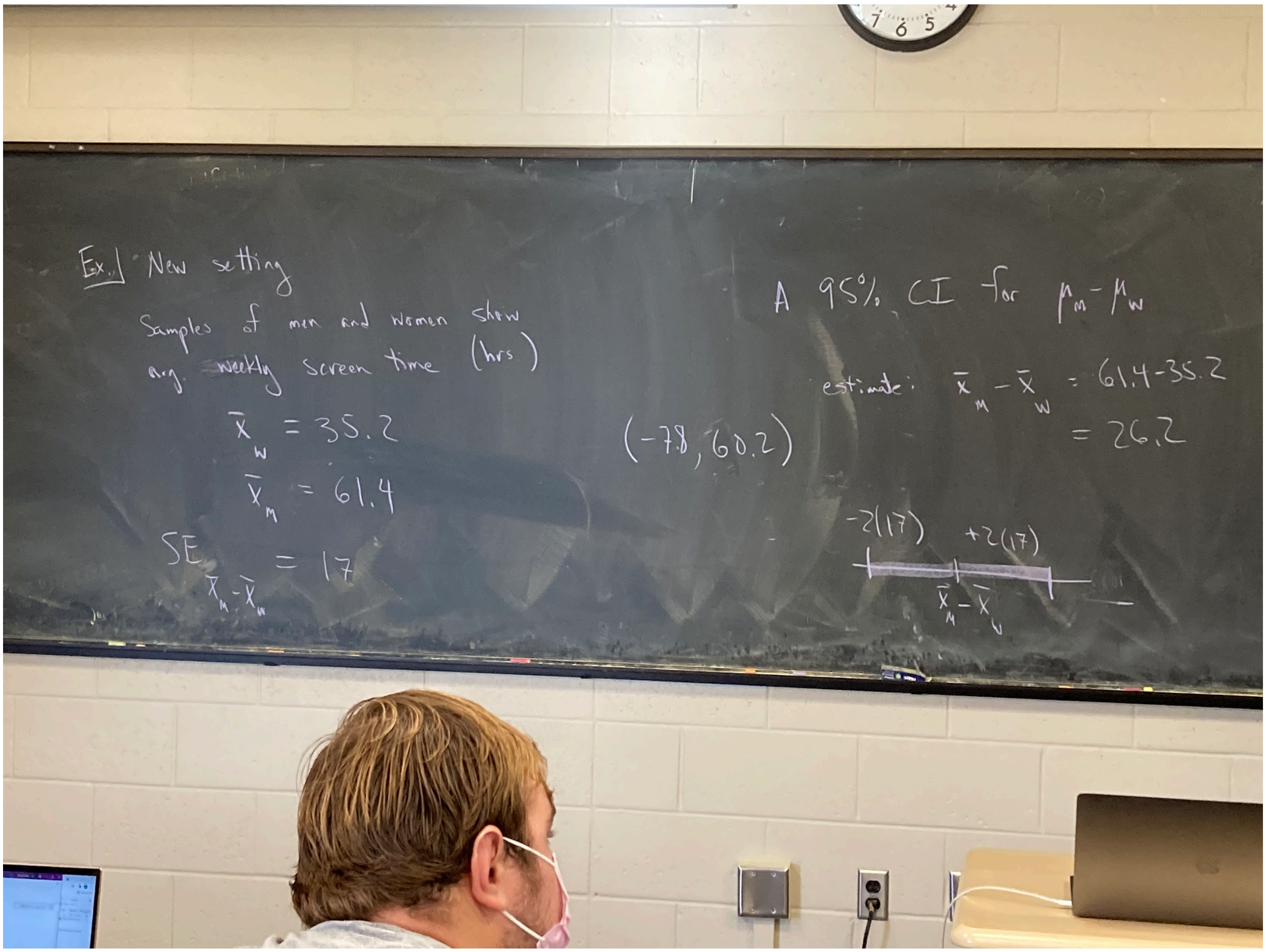
Say that $SE_{\hat{p}} = 0.016$.

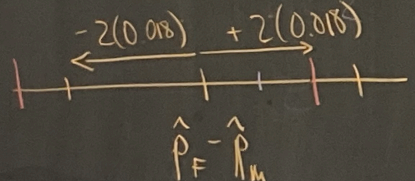
Give a 95% CI for p.

lower: $0.42 - 2(0.016)$,

upper: $0.42 + 2(0.016)$

95% CI: $(0.388, 0.452)$


Ex.) New setting
Samples of men and women show
avg. weekly screen time (hrs)
$\bar{x}_W = 35.2$
$\bar{x}_M = 61.4$
$SE_{\bar{x}_M - \bar{x}_W} = 17$
$(-7.8, 60.2)$
A 95% CI for $\mu_M - \mu_W$
estimate: $\bar{x}_M - \bar{x}_W = 61.4 - 35.2$
$= 26.2$
$-2(17)$
$+2(17)$
$\bar{x}_M - \bar{x}_W$



proportion of disagreers

Females: $\hat{p}_F = \frac{1005}{1402}$

Males $- \hat{p}_M = \frac{807}{1223}$

$\frac{1005}{1402} - \frac{807}{1223}$

95% CI for $p_F - p_M$

95% Confidence Intervals -
Estimating SE using bootstrapping

Confidence Intervals

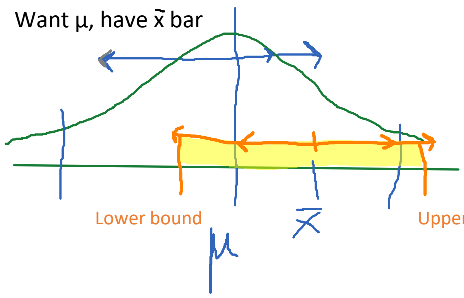
Goal: <u>Estimate:</u>

- μ population mean
- P population proportion
- ρ (rho) true correlation between two variables
- B_1 population slope
- $\mu_1 - \mu_2$
- $\rho_1 - \rho_2$

<u>Estimators:</u> Always a bit off!

- $\bar{x}$ bar
- $\hat{p}$ hat
- r
- b_1
- $\hat{x}$ hat 1 - $\hat{x}$ hat 2
- $\hat{p}$ hat 1 - $\hat{p}$ hat 2

Want μ, have $\bar{x}$ bar



Sampling dist. Normal

95% of values are within 2 SE(standard error) of center

Our first method for constructing 95% CI:

(point est.) +/- 2 SE

★ for CI=95%, double SE.

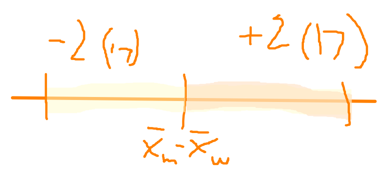
Ex. 1) Have poll saying 42% of Americans favor proposal A. (p hat)
Say that SE = 0.016
Give a 95% CI for P (population proportion)
0.42 - 2(0.016) = 0.388 = lower bound
0.42 + 2(0.016) = 0.452 = upper bound

Margin of error - the amount added or subtracted from estimate to obtain lower/upper bounds of your CI

Ex. 2) New Setting
Samples of men and women show average weekly screen time (hrs)

$X_w = 35.2$
$X_m = 61.4$

A 95% CI for $\mu_m - \mu_w$
Estimate: $X_m - X_w = 61.4 - 35.2 = 26.2$
(-7.8, 60.2)



Ex. 3) Proportion of Disagreers

Females: $\hat{P}_F = 1005/1402$
Males: $\hat{P}_M = 807/1223$
95% CI for $P_F - P_M$ = (1005/1402)-(807/1223) = 0.057